Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/AU04/001587

International filing date: 17 November 2004 (17.11.2004)

Document type: Certified copy of priority document

Document details: Country/Office: AU
Number: 2003906304
Filing date: 17 November 2003 (17.11.2003)

Date of receipt at the International Bureau: 04 January 2005 (04.01.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

PCT/AU2004/001587

Patent Office
Canberra

I, LEANNE MYNOTT, MANAGER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2003906304 for a patent by DR IAN FINDLAY as filed on 17 November 2003.

WITNESS my hand this
Sixteenth day of December 2004

LEANNE MYNOTT
MANAGER EXAMINATION SUPPORT
AND SALES

High throughput isolation and genetic analysis of prenatal samples

1 TITLE

High throughput isolation and genetic analysis of prenatal samples.

2 FIELD OF THE INVENTION

THIS INVENTION relates to the combinations of isolation/enrichment of cells from samples, the genetic identification and analysis of such cells and the automated processing of resultant data to enable high throughput genetic analysis and thus practical application. More particularly, this invention relates to high throughput enrichment of fetal cells from cervical samples, and in particular, Pap smears. In a particular form, this invention relates to use of fetal cells enriched from cervical samples for subsequent genetic analysis such as Down syndrome, sex and single gene defects. Particular embodiments of this invention utilize cell isolation procedures such as laser micro-dissection, magnetic and/or fluorescent activated cell sorting, singly and in combination, to enrich fetal cells to levels of purity that readily enable nucleic acid isolation for genetic analysis. Particular embodiments of this invention utilize techniques such as nucleic acid amplification and/or a variety of analysis techniques such as multiplex PCR to subsequently genetically analyse pooled or single cells.

3 BACKGROUND OF THE INVENTION

Genetic defects are *the* major cause of embryonic, fetal and neonatal death as well as being responsible for a large proportion of childhood disabilities. The life-long cost of these disabled children to society is enormous. Although many defects are detected by the annual ~50,000

prenatal tests in Australia, tests are only offered to high-risk mothers as they are invasive (~1% risk of miscarriage) and/or expensive. One result of current screening strategies is that the vast majority of babies with genetic defects are born to the low-risk population. Currently prenatal diagnosis of chromosomal and single gene disorders requires the withdrawal of fetal cells from the uterine cavity by invasive procedures such as amniocentesis or chorionic villus sampling (CVS). These techniques, although highly reliable, carry procedurally related risks such as miscarriage (0.5-1%), require a high level of technical expertise, can take several weeks for results and can only be performed relatively late in pregnancy. The miscarriage risk and high cost limit availability limits the availability of prenatal diagnosis to high risk mothers only.

Current prenatal risk assessment has two main limitations. Firstly, high risk mothers are offered tests even though most (~95%) do not have an affected fetus, resulting in unnecessary miscarriage risk and high cost. Secondly, affected children from low risk groups are usually not identified.

Currently prenatal diagnosis of chromosomal and single gene disorders requires the extraction of fetal cells from the uterine cavity by invasive procedures such as amniocentesis or chorionic villus sampling (CVS). These techniques, although highly reliable, carry procedurally related risks such as miscarriage (0.5-1%), require a high level of technical expertise, take several weeks for results and can only be performed relatively late in pregnancy. Thus, they are only offered to women considered at high risk due to age, genetic history or other indicative factors.

One less invasive alternative approach is to use maternal blood as a source of fetal cells for which many fetal cell enrichment methods have

been developed, for example as described in United States Patent 5,629,147, United States Patent 5,646,004 and International Publication WO 98/02528.

However major technical difficulties remain due to the extremely low numbers of fetal cells found in the maternal circulation, the extreme difficulties in isolating such cells, the positive identification of fetal cells and the presence of fetal cells from previous pregnancies which may confound identification and diagnosis.

The presence of fetal cells in the endocervical canal was first published in Shettes, 1971, Nature **230** 52. Since then there have been many studies confirming the presence of fetal cells with varying success during the first trimester (Fejgin *et al.*, 2001, Prenatal Diagnosis **21** 619; EN.REFLIST These studies all confirm that the number of fetal cells present in the endocervical canal of pregnant women is extremely low and secondly that these cells are difficult to isolate.

Accordingly, the value of maternal cervical samples as a source of fetal cells for genetic analysis has remained controversial. Additionally, there have been major concerns as to the invasiveness and safety of cervical sampling and the practicality of using cervical samples as a source of relatively low abundance fetal cells. Indeed, Overton *et al.*, 1996, J. Am. Obstet. Gynecol. **175** 382 concluded that fetal cells cannot be obtained from the endocervix by minimally invasive techniques in sufficient yield for prenatal genetic diagnosis.

More recently Cioni *et al.*, 2003 Prenatal Diagnosis **23** 168-171 confirmed that fetal cells were not detected in a consistent and reliable fashion and therefore such sampling techniques cannot be regarded as a promising tool towards minimally invasive prenatal diagnosis.

It has been shown (Kingdom et al., 1995 Obstetrics and Gynaecology, 86 pp 283-288) that fetal cells can be isolated from cervical samples such as PAP smears using fluorescently labelled antibodies to remove maternal cells (negative enrichment) and extract fetal cells (positive enrichment) with subsequent genetic analysis in a similar manner to that performed for fetal cells in maternal blood, Immunology, 30 (2-3) pp.194-201; Durrant et al., 1996 British Journal Of Obstetrics And Gynaecology, 103, (3), 219-222). However such techniques are expensive, time intensive and are limited in sample throughput to <10 samples per day, which is insufficient for the high throughput required for cost-effective clinical application. For the purposes of this invention, high throughput refers to the ability to process in excess of 50 samples per 24hr day.

Previous work has identified five main difficulties in applying performing genetic analysis from cervical cells at a high throughput level sufficient for practical application.

1. Obtaining sufficient cells. Attempts to obtain fetal cells from the cervix of pregnant women have been hampered by the need to retrieve the large number of cells required for genetic diagnosis. Although this requirement has been partially overcome by no longer requiring large numbers of cells due to recent advances such as multiplex fluorescent PCR which now allow multiple genetic analyses from single cells, single cell multiplex PCR remains a highly technical process practiced by very few laboratories worldwide due to its high complexity. Previous work has indicated that fetal cells are not detected in a consistent and reliable fashion and therefore cannot be regarded as a promising tool for prenatal analysis.

2. Isolation of fetal cells from the sample. Recent results suggest that fetal cells can be isolated and diagnosed in only ~22% of cases due to the presence of "contaminating" maternal cells. Previous approaches have generally concentrated on isolating cervical cells by morphology or cell sorting. Unfortunately, morphology grading is extremely time-consuming, expensive and generally unreliable and inaccurate. Cell sorting such as FACS or MACS has been generally unspecific resulting in either major maternal contamination and misdiagnosis or insufficient fetal cells. This remains a major limitation to practical application.
3. Genetic identification of cells to determine fetal source. It is essential to identify the isolated cells as being fetal to avoid misdiagnosis from maternal cells. Although single cell DNA fingerprinting techniques such as single cell multiplex PCR can be used to forensically identify the source of a single cell, again single cell multiplex PCR remains a highly technical process practiced by very few laboratories worldwide due to its high complexity. However results from fetal cells weré not detected in a consistent and reliable fashion and therefore cannot be regarded as a promising tool.
4. Genetic diagnosis from small cell numbers. Diagnosis from single or low numbers of cells is extremely difficult. Fetal cells have been identified in cervical samples mainly by the identification of male cells within the sample, aneuploidy screening (the primary reason for prenatal diagnosis) cannot usually be performed nor diagnosis made if the fetus is female. This requirement has been partially overcome by recent advances such as multiplex fluorescent PCR which now allow multiple genetic analyses from single cells. However again results from fetal cells were not

detected at a consistent and reliable enough level to be considered as a promising tool towards minimally invasive prenatal diagnosis.

5. Sample collection. The most recent work on PAP smears have utilised so called "thick section" PAP smears. However clinical PAP smears collections are moving towards a "thin section" smears which have not previously been used for fetal cell isolation.

Each of the five major difficulties above have singly, and particularly in combination, prevented practical application of non-invasive prenatal diagnosis.

However, by improving current methods and developing new methods in combination, we have invented a novel high throughput method, to rapidly and efficiently accurately target and genetically diagnose affected babies in both high-risk mothers and previously designated low risk pregnancies. For the first time non-invasive prenatal diagnosis is now a practical application and will have enormous health, social and economic benefit.

4 SUMMARY OF THE INVENTION

Notwithstanding the prior art teaching that cervical samples are very poor sources of fetal cells for genetic analysis, the present inventors have developed a high throughput method encompassing reliable and efficient combination of methods which isolate and enrich fetal cells from cervical samples, performed genetic analysis (DNA fingerprinting and/or genetic diagnosis/screening) with subsequent data analysis which enables sufficient high-throughput processing for practical application.

In one aspect, the invention provides a high throughput method of cell isolation including the step of enriching one or more fetal cells from a cervical sample.

In another aspect, the invention provides a high throughput method of obtaining a nucleic acid sample, including the step of isolating a nucleic acid from one or more fetal cells that have been enriched from a cervical sample.

In yet another aspect, the invention provides a high throughput method of genetic analysis including the step of analyzing a nucleic acid obtained from one or more fetal cells that have been enriched from a cervical sample.

In a further aspect, the invention relates to the high throughput use of one or more fetal cells enriched from a cervical sample for genetic analysis.

In a further aspect, the invention combines incremental improvements over a number of steps (isolation, DNA fingerprinting and genetic analysis and data processing), which combine to give a significantly increased throughput sufficient for practical application.

In a further aspect, the invention comprises multiple aspects above.

It will also be appreciated that other analyses also contemplated by the present invention include biochemical analysis, morphological analysis, histology, cytology, cell culture and the like as well as a variety genetic analysis including nucleic acid amplification methods such as PCR, CGH (comparative genome hybridization), SNPs (single nucleotide polymorphisms), FISH (fluorescent in situ hybridization) and the like.

In a still further aspect, the invention relates to the high throughput use of one or more fetal cells enriched from a cervical sample for the isolation of a nucleic acid sample.

In a yet further aspect, the invention relates to high throughout use of a cervical sample for enrichment of one or more fetal cells for the isolation of a nucleic acid sample.

In a still yet further aspect, the invention relates to high throughout use of a cervical sample for enrichment of one or more fetal cells for genetic analysis.

The invention also relates to high throughout enrichment steps described herein to enrich fetal material from cervical samples.

The invention also relates to the high throughout automation of steps described herein to enrich fetal material from cervical samples.

The invention also relates to any combination of isolation/enrichment techniques with any combination of nucleic amplification and/or genetic identification techniques such as DNA identification, and/or genetic diagnosis and/or automated data analysis to allow sufficiently high throughput for practical application.

It will be appreciated by the skilled addressee that any and all of these high throughput techniques can be readily applied to a variety of samples including, but not limited to, blood, vaginal cells, PAP smears (both thin and thick specimens).

Preferably, a cervical sample is obtained using an endocervical brush or cytobrush.

More preferably, the cervical sample is a Pap smear.

In a preferred embodiment, the invention provides a method of fetal cell analysis including any combination of the steps of:

(i) enriching fetal cells from a Pap smear sample according to physical characteristics such as size, morphology and/or granularity; and/or

(ii) positively selecting fetal cells from the cells enriched in step (i) and/or in using at least one antibody that binds a fetal cell antigen.

(iii) Amplification of generic nucleic acid from isolated sample from step (ii)

(iv) Genetic identification product from step (iii) using techniques such as DNA fingerprinting

(v) Genetic analysis of product from step (ii), (iii) and/or (iv) including but not limited to specific genetic analysis methods such as multiplex PCR, SNPs, CGH, FISH, RT-PCT and the like. It will be appreciated that steps (v) and (vi) can be combined into a single analysis procedure

(vi) Detection of products from steps (iii), (iv) and/or (v) for example utilising nucleic acid separation technologies such as a DNA sequencer

(vii) automated data processing to create analysis report from step (vi)

Throughout this specification, unless otherwise indicated, "comprise", "comprises" and "comprising" are used inclusively rather than exclusively, so that a stated integer or group of integers may include one or more other non-stated integers or groups of integers.

5 BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figure 1. Diagram of Invention Overview

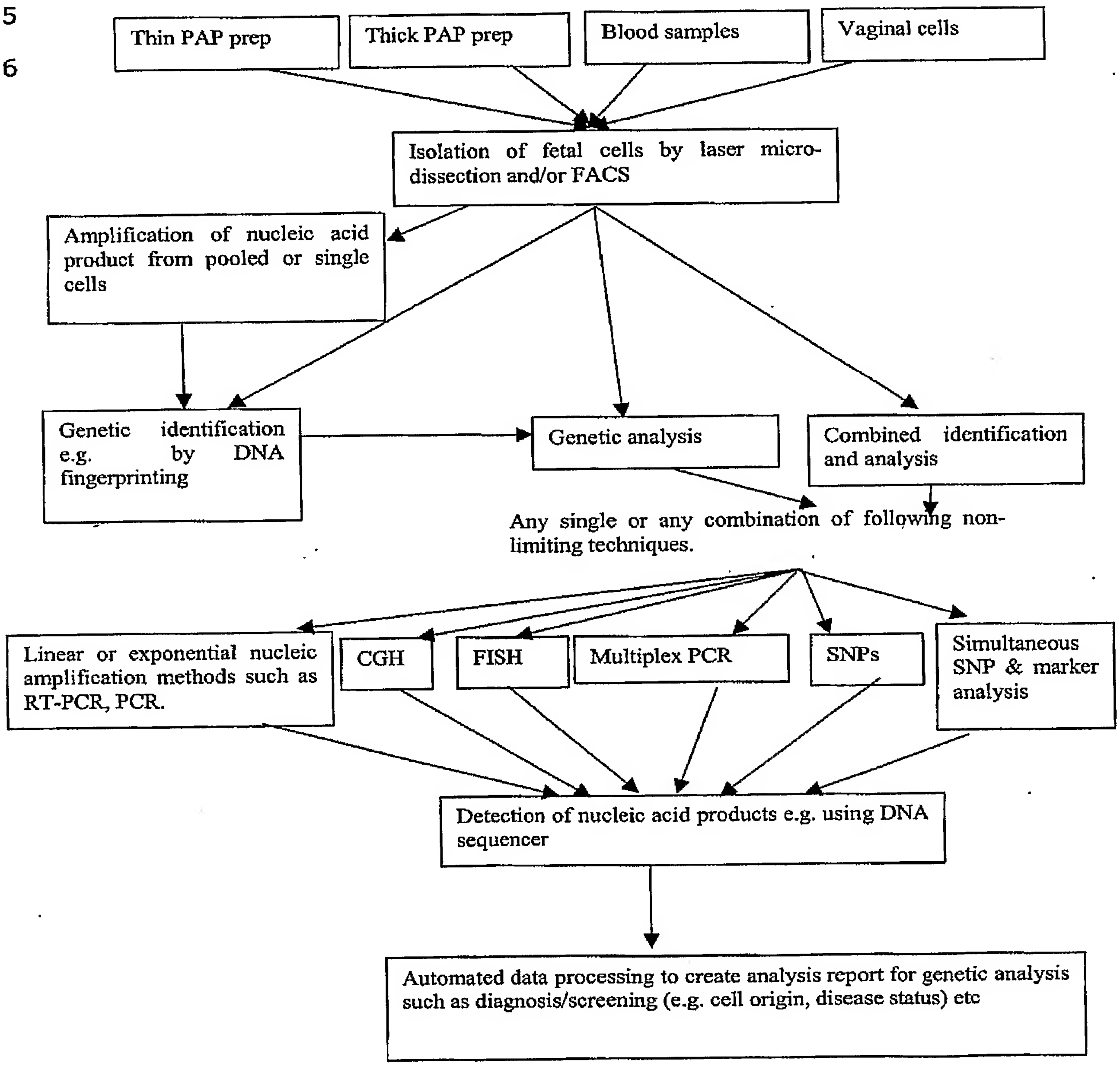

Diagram is for simplicity only and should not be construed to indicate or limit possible embodiments.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a variety of methods applicable to genetic analysis of fetal cells from maternal samples, and in particular, from cervical Pap smears. Such methods include the steps detailed previously comprising cell isolation, nucleic acid amplification, genetic identification and analysis and automated data processing.

It will also be appreciated that the present invention is applicable to isolation or enrichment of other cells of non-maternal origin including, but not limited to, embryonic cells, sperm cells and any cells of cytotrophoblast or syncytiotrophoblast origin.

It should also be appreciated that the present invention is applicable to isolation or enrichment of other cells of non-maternal origin from a variety of other sources such as maternal blood, vaginal cells and the like.

For the purposes of this invention, by "*isolated*" is meant material that has been removed from its natural state or otherwise been subjected to human manipulation. Isolated material may be substantially or essentially free from components that normally accompany it in its natural state, or may be manipulated so as to be in an artificial state together with components that normally accompany it in its natural state.

By "*enrich*", "*enriched*" and "*enrichment*" in the context of cell isolation is meant that cells are obtained in a higher frequency of proportion compared to their frequency or proportion in a starting sample prior to enrichment. In this context enrichment is also taken to include 100% enrichment where the fetal cell or cells exist in the absence of maternal cells.

Suitably, fetal cells are enriched from a cervical sample. Such samples include and encompass any sample obtained from the endocervix

inclusive of endocervical lavage, aspiration, swabbing; cytobrush samples; transcervical samples (TCCs) and Pap smears

Preferably, the cervical sample is a Pap smear.

As used herein, a Pap smear is a biological sample comprising one or more cells collected, obtained as a scraping from the cervix.

Typically, a metal or plastic instrument such as a speculum is placed in the vagina to allow visualization of the interior of the vagina and the cervix. A sampling instrument such as a small wooden spatula is used to scrape the outside of the cervix and thereby obtain the cervical sample.

For the purposes of cervical cancer screening, the scrapings are placed on a glass slide and used for microscopic examination to detect changes in the cells of the cervix. Pap smears are a routine and safe screening procedure to find early warning signs of cervical cancer. The present invention provides a new use of Pap smears as a source of fetal cells for enrichment and subsequent analysis.

In the context of the present invention, said one or more cells typically comprises maternal cells and fetal cells.

For the particular purpose of fetal cell isolation, it is preferred that the Pap smear is obtained at between 5 and 31 weeks gestation.

Cell enrichment may be performed by one or more high throughput cell isolation methods including complement-mediated lysis, flow cytometry, magnetic bead separation, panning, charge flow separation, laser microdissection and cell culture methods that promote selective propagation of cells to be enriched.

Each cell enrichment method may be performed alone or in combination with one or more other methods to thereby achieve a desired level of cell enrichment or purity.

Additional treatments may be utilized that facilitate cell isolation and enrichment, for example in one embodiment protease treatment (*e.g.* trypsin digestion) of cervical samples may be performed prior to density gradient enrichment.

5.1 Cell enrichment/isolation

It will be appreciated that fetal cell enrichment may be achieved using physical characteristics such as size, granularity and/or using antibodies directed to fetal antigens not expressed, or expressed at low levels, by maternal cells.

Alternatively, fetal cells may be enriched by virtue of their non-expression of maternal or non-fetal antigens.

Accordingly, fetal cell enrichment may be performed by negative depletion of maternal cells and/or positive selection of fetal cells according to antigen expression.

Antigens that may be applicable to antibody-based enrichment include, but are not limited to, CD71, γ globin (fetal) and ζ globin (embryonic), glycophorin A, CD36, Fkl-1, EPO-R, CDw50, CD45, human chorionic gonadotrophin (HCG), placental alkaline phosphatase, human placental lactogen (FD0202N), folate binding protein (LK26) and HLA antigens such as HLA-Class II, for each of which specific antibodies are readily available.

Preferred antigens are human placental lactogen (FD0202N) and folate binding protein (LK26).

In the broadest sense, antibody-based enrichment may utilize any technique that selects cells (*i.e* positive selection) or depletes cells (*i.e* negative selection) according to antigen expression or non-expression, as

the case may be. A non-exhaustive list includes panning, complement-mediated lysis, fluorescence-activated cell sorting (FACS) and magnetic activated cell sorting (MACS).

It will also be appreciated that the aforementioned techniques may be used alone or in sequential combination to enrich fetal cells.

Preferred methods utilize FACS. FACS can enrich samples by using physical cellular characteristics including but not limited to size, shape, granularity and the like and/or fluorescently labelled antibodies.

For FACS enrichment, fluorescently-labeled antibodies are bound to the cells of interest. These cells are then passed through the excitation laser in a single cell stream and measured for size, granularity and fluorescent activity. Specific parameters are set and cells that fall within those parameters (*e.g.* fluorescence, forward light scatter, side scatter) are collected by a cell sorter into receptacles such as 96 or 384 well plates to facilitate high throughput.

In a preferred embodiment, fetal cells are enriched by FACS using antibody to placental lactogen, trypsin release of bound beads followed by FACS enrichment using antibody to folate binding protein (LK26).

Charge Flow Separation

Charge flow separation uses dielectrophoretic forces which occur on cells when a non-uniform electrical field interacts with field-induced electrical polarization. Depending on the dielectric properties of the cells relative to their suspending medium, these forces can be positive or negative, directing the cells toward strong or weak electrical field regions. Because cells of different types or in distinct biological states have different dielectric properties, differential dielectrophoretic forces can be applied to

drive their separation into purified cell populations (Wang *et al.*, 2000. Analytical Chemistry **72** 832-839).

Fetal cells may be enriched by selective growth in the presence of appropriate cytokines and culture conditions that favor the selective proliferation of fetal progenitor cells over maternal cells. Selective growth may be performed after initial isolation or enrichment by one or more other enrichment methods.

For example, fetal nRBC's (nucleated red blood cells) may be cultured after gradient enrichment and/or MACS enrichment in culture media containing many fetal NRBC growth factors (Bohmer *et al.*, 1998, Br J Haematol **103** 351-360). It is also contemplated that culture with fetal NRBC growth factors may stimulate a much higher basal proliferative capacity than mature progenitor cells and that this can be enhanced by addition of cytokine cocktails such as flt-3 ligand and thrombopoetin (Holzgreve *et al.*, 2000, Baillieres Best Pract Res Clin Obstet Gynaecol **14** 709-722).

In light of the foregoing, a preferred embodiment of the invention provides a method of fetal cell isolation and analysis including the steps of:

(i) improved high throughput enriching of fetal cells from a Pap smear sample according to physical and/or fluorescent characteristics where at least one antibody binds to a fetal cell antigen.

Preferably, step (i) includes the sequential steps of:

(a) protein digest to release cellular mixture into discrete cells

(b) using FACS to determine and separate fetal cells by physical characteristics

(c) using FACS to determine and separate fetal cells based on fluorescent characteristics using antibodies such as placental lactogen;

(ii) Improved high throughput amplification of nucleic acid product from pooled or single cells using a generic kit such as Genomiphi.

(iii) Genetic identification of sample using DNA fingerprinting. Genetic identification can be undertaken using STR profiling as previously published (Findlay *et al.*, 1997, Nature **389** 355-356)

(iv) Genetic analysis using a variety of non-limiting techniques such as RT-PCR, CGH, FISH, multiplex PCR, SNP analysis, and/or simultaneous analysis of any combination of above.

Preferably, step (iv) includes the sequential steps of:

(a) High throughput genetic analysis

(b) Detection of genetic product

Improved high throughput automated data processing to create analysis report. Analysis reports can include indicators of diagnosis/screening markers disease status as well as factors such as cell origin. Automated analysis provides the capacity to analyse many millions of analyses parameters extremely quickly and thus provide high throughput analysis.

The combination of each of the incremental improvements in the above steps allow high throughput processing and analysis of fetal samples

Other uses of enriched cells is for subsequent genetic analysis, biochemical analysis, morphological analysis, histology, cytology, cell culture and the like.

5.2 Genetic analysis

A more preferred use is for genetic analysis.

As used herein, "*genetic analysis*" and "*genetic diagnosis*" are used interchangeably and broadly cover detection, analysis, identification and/or characterization of isolated genetic material and includes and encompasses terms such as, but not limited to, genetic identification, genetic diagnosis, genetic screening, genotyping and DNA fingerprinting (also commonly known as STR profiling) which are variously used throughout this specification.

The term "*nucleic acid*" as used herein designates single-or double-stranded mRNA, RNA, cRNA, RNAi and DNA inclusive of cDNA, genomic DNA and DNA-RNA hybrids.

A "*polynucleotide*" is a nucleic acid having eighty (80) or more contiguous nucleotides, while an "*oligonucleotide*" has less than eighty (80) contiguous nucleotides.

A "SNP" is a single nucleotide polymorphism.

A "*primer*" is usually a single-stranded oligonucleotide, preferably having 12-50 contiguous nucleotides which, for example, is capable of annealing to a complementary nucleic acid "template" and being extended in a template-dependent fashion by the action of a DNA polymerase such as *Taq* polymerase, RNA-dependent DNA polymerase or Sequenase™.

By "*genetic marker*" or "*marker*" is meant any locus or region of a genome. The genetic marker may be a coding or non-coding region of a genome. For example, genetic markers may be coding regions of genes,

non-coding regions of genes such as introns or promoters, or intervening sequences between genes such as those that include polymorphisms (such as single nucleotide polymorphisms), tandem repeat sequences, for example satellites, microsatellites, short tandem repeats (STRs) and minisatellites, although without limitation thereto.

A "*probe*" may be a single or double-stranded oligonucleotide or polynucleotide, suitably labeled for the purpose of detecting complementary sequences in Northern or Southern blotting, for example.

Genetic analysis may be performed by any method including, but not limited to, fluorescence *in situ* hybridization (FISH), primed *in situ* synthesis (PRINS) and nucleic acid sequence amplification, preferably in the form of multiplex fluorescent PCR amplification (MFPCR).

Examples of fluorescent in *situ* hybridization (FISH) and Primed In Situ Synthesis (PRINS) may be found in Findlay *et al.*, 1998, J. Assisted Reproduction & Genetics **15** 257.

5.2.1 Multiplex fluorescent PCR

As used herein, "*multiplex amplification*" or "*multiplex PCR*" refers to amplification of a plurality of genetic markers in a single amplification reaction.

MFPCR has been shown to be a reliable and accurate method for determining sex (Salido *et al.*, 1992, Am. J Human genetics **50** 303; Findlay *et al.*, 1994a, Human Reproduction, **9** 23; Findlay et al., 1994b, Advances in Gene Technology: Molecular Biology and Human Genetic Disease. Vol 5, page 62. Findlay *et al.*, 1995, Human Reproduction **10** 1005-1013; Findlay *et al.*, 1998c, *supra)* diagnosing genetic diseases such as cystic fibrosis (Findlay *et al.*, 1995, *supra*), detecting chromosomal aneuploidies and in genetic analyses for genetic identification, such as typically referred to as DNA fingerprinting (Findlay *et al.*, 1997, Nature **389** 355-356).

With regard to genetic markers for genetic analysis, preferred genetic markers are STR and/or SNP markers. International Application PCT/AU02/01388 provides an extensive array of STR markers and primers together with MFPCR methodology to successfully amplify multiple STR markers from limiting amounts of nucleic acid template.

Although from the foregoing a preferred method of genetic analysis is PCR, nucleic acid sequence amplification is not limited to PCR.

Nucleic acid amplification techniques are well known to the skilled addressee, and also include ligase chain reaction (LCR) as for example described in Chapter 15 of Ausubel *et al.* CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley & Sons NY, 1995-1999); strand displacement amplification (SDA) as for example described in U.S. Patent No 5,422,252; rolling circle replication (RCR) as for example described in Liu *et al.*, 1996, J. Am. Chem. Soc. **118** 1587 and International application WO 92/01813 and by Lizardi *et al.*, in International Application WO 97/19193; nucleic acid sequence-based amplification (NASBA) as for example described by Sooknanan *et al.*,1994, Biotechniques **17** 1077; and

Q-β replicase amplification as for example described by Tyagi *et al.*, 1996, Proc. Natl. Acad. Sci. USA **93** 5395.

The abovementioned are examples of nucleic acid sequence amplification techniques but are not presented as an exhaustive list of techniques. Persons skilled in the art will be well aware of a variety of other applicable techniques as well as variations and modifications to the techniques described herein.

As used herein, an "*amplification product*" refers to a nucleic acid product generated by a nucleic acid amplification technique.

Although the invention also contemplates use of nucleic acid other than DNA, preferably the nucleic acid is DNA.

More preferably, the nucleic acid is genomic DNA.

Isolation of cellular nucleic acids is well known in the art, although the skilled person is referred to Chapters 2, 3 and 4 of Ausubel *et al.* CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley & Sons NY, 1995-1999), for examples of nucleic acid isolation.

5.2.2 SNP analysis

Single Nucleotide Polymorphisms (SNP) are the most frequent form of variation found in the genome, estimated to occur every 1000 bases. SNP genotyping has multiple applications such as predictive medicine, personal medicine, forensic identification and pharmacogenomics. SNP genotyping has already been used to investigate a number of disorders such as cystic fibrosis, Factor V Leiden mutation, and factors such as A, B, O and Rh blood grouping. However conventional SNP analysis is limited by the relatively high amount of extracted DNA usually required (up to 100ng) for analysis. However in genomic analysis, there is increasing demand to both maximize data by performing multiple analyses and secondly to analyze minimum

amounts of sample, even to the single cell level. Although multiple SNP analyses can be performed routinely, the degree if sensitivity is still far from single cell level analysis. Multiplex single cell SNP analysis has been problematic and again is not amenable to the high throughput processing required of clinical application.

Preferred sources of nucleic acids are mammals, preferably humans.

The invention also contemplates genetic analysis of non-human samples such as from cows, sheep, horses, pigs and any other mammal including companion animals, sporting animals and livestock, although without limitation thereto.

So that the invention may be readily understood and put into practical effect, reference is made to the following non-limiting examples.

6 MATERIALS AND METHODS

6.1 Fetal cell Isolation

Informed consent was obtained from pregnant women between 7 and 31 weeks gestation. A pap smear cervix brush (Rovers Medical Devices, Lekstraat, The Netherlands) was inserted through the external os to a maximum depth of two centimeters. The brush was then removed whilst rotating a full turn. The material that was caught on the brush was smeared on a slide which was used for routine cervical cancer screening. The remaining material on the brush was included in the study. Some pregnant women were undergoing pregnancy termination immediately after the pap smear sample was taken, and in these cases the pap smear sample was not used for routine cervical screening and the entire sample was available for analysis.

Cells were washed off the cervix brush into Dulbeccos PBS (Invitrogen, Melbourne Australia). The cells were then spun at 402g in a Sigma 4K15 centrifuge (Sigma, St Louis, USA), the supernatant was poured off and the remaining material transferred to a 1.5mL centrifuge tube. The cells were then spun at 3000rpm in a biofuge pico (Kendro, Ashville, North Carolina), the supernatant was then poured off and the cells were resuspended in 800uL PBS. A buccal swab was also taken from the mother to provide an uncontaminated source of maternal cells.

6.2 Trypsin Digestion

- 3 X 200uL + 1 X 50uL cell suspension from each sample
- 200uL (or 50uL) 2.5% Gibco Trypsin/EDTA in PBS was added
- Incubated 37 degrees for 2 hours
- 1mL PBS was added and then centrifuged 3000rpm 5 minutes (Biofuge pico)
- Supernatant was removed

50uL initial material sample was placed in –20°C freezer for later analysis

6.3 Isolation of fetal cells by physical characteristics using FACS

1. Cells should be washed 2-3x with FACS buffer (PBS supplemented with either 1% BSA or 5% FBS and containing 0.05% NaN_3).
2. Suspend the cell pellet from the final wash in 50 microliters FACS buffer (or more if more than one analysis is to be done on a single sample).
3. Incubate for 30 minutes on ice.

4. Wash cells 2-3x with FACS buffer and suspend in 200-300 microliters FACS buffer for analysis.

Fetal cells may be isolated by any of the aforementioned cell isolation methods.

In all cases samples from non-pregnant women are run as control cases to determine the base-line level of non-specificity.

Preferably, said one or more fetal cells are isolated by FACS sorting.

Said one or more fetal cells may be isolated from any pregnant mammal.

Preferably, said one or more fetal cells are isolated from a pregnant human.

When a fetus is at increased risk for genetic defects such as chromosomal anomalies, prior art prenatal diagnosis is by invasive procedures such as either chorionic villus sampling (CVS) in the late 1st trimester or amniocentesis in the 2nd trimester of pregnancy. By the third trimester, a combination of CVS and amniocentesis, or even fetal blood sampling, may be necessary.

A rapid, less-invasive and low cost method of prenatal diagnosis involves genetic diagnosis from fetal cells shed into the cervical sump at 6-20 weeks of gestation. These samples are obtained from the cervix by cytobrush in a manner identical to a PAP smear, which is similar to but significantly less invasive than invasive transcervical sampling.

Although promising, previous work has identified several major difficulties. Firstly the need to obtain the large numbers of fetal cells required for genetic diagnosis. Secondly the isolation of fetal cells from the cervical sample is extremely difficult as recent results suggest that fetal cells could be isolated and diagnosed in only ~22% of cases due to the presence of "contaminating" maternal cells. Previous approaches have generally concentrated on isolating fetal cells by morphology or cell sorting. Unfortunately, morphology grading is extremely time-consuming, expensive and generally unreliable and inaccurate. Alternative cell sorting techniques involves antibody-labelled slides to capture fetal cells, which is generally unspecific resulting in major maternal contamination and misdiagnosis or insufficient fetal cells. Thirdly the difficulty in positively identifying the isolated cell as being fetal rather than maternal; previous approaches have determined male fetal cells which identifies the fetus as the mother is female but female signals could either indicate female cells or maternal contamination and thus misdiagnosis.

Fourthly the difficulty of obtaining genetic diagnosis from small cell numbers. Although fetal cells have been identified in cervical samples (mainly by identifying male cells within the sample) aneuploidy screening (the primary reason for prenatal diagnosis) cannot usually be performed nor diagnosis made if the fetus is female due to the risk of contamination causing misdiagnosis.

Finally, although PAP sampling collection is theoretically much safer than CVS and amniocentesis as PAP smears have been taken during pregnancy for many years, relative safety remains to be fully evaluated.

According to the present invention, it is preferred that said fetal cells are present in a maternal uterine cavity or endocervical canal sample, particularly a transcervical sample. Methods of isolating fetal cells include cervical cotton swab, cytobrush, aspiration of cervical mucus, lavage of the endocervical canal and uterine lavage. Samples can be obtained from transcervical aspiration of mucus from just above the internal os or the lower uterine cavity. Lavage is generally conducted with a saline wash, but other isotonic solutions are suitable. Typically, endocervical lavage with 5-10ml or intrauterine lavage with 10-20ml saline provides sufficient fetal cells upon separation from maternal cells. The sample may be collected using a combination of methods.

Preferably, cell samples are isolated from a female human in the first trimester of pregnancy or when the fetus is between 6 to 17 weeks gestation. The sample can be in any solution which maintains cell integrity and minimizes cell lysis or damage, preferably a physiological solution, or more preferably, a saline solution or tissue culture medium with or without the addition of sera.. The sample is preferably stored at 0°C to 4°C until use to minimize the number of dead cells, cell debris and cell clumps.

Preferably, to aid fetal cell separation, clumps of cells are preferably treated to obtain a suspension of single cells. The clumps may be separated by techniques known to a skilled person, such as enzymatic, chemical or mechanical separation. For example, enzymatic separation may utilise protease or trypsin. Chemical separation may utilise acetyl cysteine and mechanical separation may involve gentle teasing, aspiration or micromanipulation.

The number of fetal cells in the sample varies depending on factors including the age of the fetus, method of sampling, number and frequency

of samplings, the vigour of sampling and the volume aspirated.

Maternal uterine cavity or endocervical canal samples typically contain at least two main types of nucleated fetal cells: cytotrophoblasts and syncytiotrophoblasts cells.

Fetal cells can be isolated either by selecting fetal cells from maternal cells (positive selection) or isolating the maternal cells from the fetal cells (negative selection) or most preferably a combination of both. Preferably, the nucleated fetal cells are retained in the purified sample.

Preferably, the maternal cells are labelled with an antibody for a selected maternal antigen. In that case, fifty percent of the time, the sample will contain unlabelled fetal cells. Preferably, the maternal cells are double-labelled. When the selected maternal antigens are encoded by alleles of different genetic loci, three out of four times, fetal cells are unlabelled or single-labelled depending on whether the fetus inherits one or neither of the maternal alleles. When the selected maternal antigens are encoded by alleles of the same genetic locus, fetal cells are single-labelled. The separation procedure is described below using double-labelled maternal cells.

Suspension media, use of protein supplements such as 5-10% BSA or HSA and appropriate cell concentrations for FACS-based separation are well known in the art and described in references such as Practical Flow Cytometry *supra* and Current Protocols in Immunology Eds Coligan *et al.*, *supra*.

More preferred as a protein supplement is about 5% autologous plasma, which can be harvested from a purified blood sample and is non-immunogenic.

The cells of the blood sample, preferably purified cells, are labelled with fluorescent antibodies specific for the antigens encoded by at least one maternal polymorphic locus, selected as described previously. The antibodies can be polyclonal or monoclonal, preferably monoclonal. Preparation of polyclonal and monoclonal antibodies for an antigen of interest is well known. Furthermore, there is a vast supply of potentially useful antibodies, such as to human HLA antigens, that are commercially available or available from hybridoma depositories such as the ATCC.

For example, in the case of separation using HLA antigens, HLA antigen-specific antibodies are commercially available. Typically the HLA Class 1 loci (A, B and C) and the Class 11 DR and DQ loci are determined by serological methods. Therefore, antibodies specific for those antigens are readily available. Sources of HLA antigen-specific antibodies include Genetic Systems (Seattle, Wash.) and C6 Diagnostics (Mequon, Wis.). Blood group antigens are also determined serologically and the antibodies are commercially available.

The antibody is labelled with a dye that facilitates, cell sorting, particularly a fluorochrome, Suitable dyes for FACS analysis and/or separation are well known in the art. Examples of dyes are described in Practical Flow Cytometry (Second Edition), *supra*, at pages 11 5-198 and in Chapter 5 of Current Protocols in Immunology, *supra*. Preferred dyes are fluorochromes including fluorescein (e.g., fluorescein isothiocyanate--F1TC), rhodamine (e.g., tetramethylrhodamine isothiocyanate--TRITC), phycoerythrin (PE), allophycocyanin (APC) and Texas Red (Molecular Probes, Eugene, Oreg.).

For four-color flow cytometric sorting, cells can be labelled with antibodies for antigens expressed by four alleles. In that case, preferably,

the antibodies are specific for both antigens expressed by the alleles of two maternal HLA loci. Maternal cells are labelled with all four fluorochromes. Fetal cells are labelled with two of the four fluorochrornes when none of the nontransmitted maternal alleles is inherited from the father. By using four fluorochromes from two loci, the fetal cells remain distinguishable from the maternal cells even when the fetus inherits one of the nontransmitted maternal alleles from the father. A second staining is only necessary when the fetus inherits both nontransmitted maternal alleles from the father. When the antibodies are for antigens expressed by three or four maternal loci, using the additional dyes increases the likelihood that the fetus did not inherit each of the maternal alleles.

Fetal cells are only indistinguishable from maternal cells by the method of the present invention in the case where the fetus inherits all six non-transrnitted maternal alleles from the father.

For antibody labeling, cells are preferably incubated at about 4°C to maintain cell integrity. Incubation for about 30 minutes at 4^{0}C is usually sufficient for substantially complete antibody binding. The sample is preferably mixed, as by using a hematology blood rocking device, during the incubation to ensure contact of the antibodies with the cells. Preferably, the incubation is performed in the dark when using a fluorochrome label. Secondary reactions (e.g. incubation of fluorochrome-labelled avidin with biotin labelled cells) are performed in the same manner.

In a preferred embodiment, an additional selection criterion is DNA content. Fetal cells having greater than 2C DNA content can be determined using a number of vital-staining fluorochromes such as the Hoechst dyes, DAP1 (4-6-diamidino-2-phenylindole), hydroethidine and 7-aminoactinomycin D (7AMD). The fluorochrome used depends on the labels

used to select the fetal cells. A second laser capable of emitting UV light is required to excite Hoechst and DAPI dyes. Each of the above-described dyes can be used with F1TC and PE.

The ability of the cell sorter to separate maternal and fetal cells ultimately depends on the percentage of fetal cells in the sample. To obtain a fetal cell sample that is at least about 60% pure (60% of the sorted cells are fetal cells), the fetal cells should constitute about 0.001% of the maternal cells or greater. Preferably, the sample contains 80%, more preferably 90% fetal cells post-sorting.

When 100% purity is desired, the sorted cells can be plated for subsequent analysis. For example, cell suspensions containing an individual cell can be isolated within a preselected volume of suspension medium by limiting dilution. Drops containing individual cells can then be placed in suitable pre-made containers (e.g. 96 well plates) for subsequent nucleic acid amplification and/or analysis.

For PCR analysis, analysis can be performed using a single, unambiguously identified fetal cell.

Alternatively, ways can be envisaged of identifying monozygosity (indicative of the presence of a monogenic disease) in a mixed cell population containing minimal fetal material including as few as one fetal cell in ten cells. Following sorting, the separated cells can be washed twice in a physiologic buffer and resuspended in an appropriate medium for any subsequent analysis to be performed on the cells.

Following the present recovery method, whether based on solid phase (eg. magnetic beads or panning) or FACS separation, the fetal cells can be used in the same manner as fetal cells obtained by other methods such as anmiocentesis and chorionic villus biopsy. The cells can be used as a source

of DNA for analysis of the fetal alleles, as by polymerase chain amplification. PCR analysis methods may be used to detect, for example, fetal sex, beta - thalassemia, phenylketonuria (PKU), and Duchennes muscular dystrophy.

Alternatively, the cells can be cultured in the same manner as biopsy materials for karyotyping analyses. However, the incubation period may be significantly shortened if a DNA content of greater than or equal to 2C is used as a selection criterion.

6.4 Isolation of antibody labelled fetal cells by FACS

Samples were placed through the following protocol:

1. 10 minutes in 10% Donkey serum (Sigma-Aldrich, St. Louis, USA) diluted in PBS
2. Donkey serum was removed
3. 100uL of each of the following Primary Antibodies, 1/100 dilution in 10% donkey serum (*i.e.* final dilution 1/200) was then added.
4. ab7816 Rabbit anti-human chorionic gonadotrophin antibody (Abcam, Cambridge, UK) and NCL-PLAP Mouse anti- placental alkaline phosphatase antibody (Novocastra Laboratories, Newcastle, UK).
5. Samples were then incubate 90 minutes at room temperature in humidified chamber
6. Samples were then washed 3 X 5 minutes in PBS
7. 100uL of each following Secondary antibody 1/200 dilution in PBS (ie final dilution 1/400) was then added.

8. Donkey FITC polyclonal to rabbit IgG (Abcam, Cambridge, UK) and Chicken Rhodamine polyclonal to mouse IgG (Abcam, Cambridge, UK).
9. Samples were then incubated 45 minutes at room temperature in humidified chamber
10. Samples were then washed 2 X 5 minutes PBS

In all cases samples from non-pregnant women are run as control cases to determine the base-line level of non-specificity.

To improve the accuracy, reliability and cost effectiveness of non-invasive prenatal genetic diagnosis from pap smears, it is necessary to develop enrichment strategies that both reduce the concentration of contaminating cells and secondly recover as many fetal cells as possible. The use of digesting enzymes such as trypsin or collagenase increase the number of fetal cells retrieved.

In one embodiment, genetic analysis of fetal cells isolated from pap smears preferably requires a number of serial enrichment strategies in order to provide a reliable source of relatively uncontaminated fetal cells. Initial FACS enrichment strategies usually identify cells using physical characteristics such as density, charge or size. Although single cycles are not highly specific they do reduce target cell loss and are relatively low in cost, therefore multiple cycles, either on physical or fluorescent characteristics, are utilised to maximise specificity whilst maintaining cellular yield. Secondary enrichment strategies such as antibody staining will often identify cells using specific cellular traits. Primary and secondary enrichment strategies must work in unison to provide a reliable source of uncontaminated fetal cells yet achieve maximum yield. Performing multiple

cycles to improve purity and/or yield are not significant time or sample limiting steps.

6.5 Cell processing

Cells were isolated by the FACS machine into 96 well plates containing 1uL lysis buffer (200mM potassium hydroxide/50mM Dithiothretol), and incubated at 65°C for 10 minutes. 1uL neutralising buffer (300mM KCl/900mM Tris-HCl pH8.3/200mM HCl) [Cui, 1989] was then added. Cells were stored at –80°C until MFPCR.

6.6 Genetic identification using DNA fingerprinting

Multiplex fluorescent PCR of Amelogenin and STRs listed in Table 1 or more preferably listed in Table 3, D3S1358, D5S818, D7S820 and CSF1P0 THO, D21S11, D18S51, VWA, FGA, D3S1358, D5S818, D7S820, CSF, TPOX was performed on isolated cells. Each reaction contained forward and reverse primers, 1 X PCR buffer (Applied Biosystems, USA), 1.5mM $MgCl_2$ (Applied Biosystems, USA), 1.25mM each dNTP (Gibco, Life Technologies, Melbourne, Australia) and 1 unit HotStart Taq (Qiagen, Australia). PCR conditions were 94°C for 2 minute denaturation followed by 45 cycles of 94°C/10 second denaturation, 57°C/1 minute annealing and 68°C/30 second extension.

PCR product was processed using Ammonium acetate/Ethanol Clean-up. Post clean-up processing involved adding 2uL of cleaned-up product to 3uL loading buffer (Amersham Biosciences, Piscataway, New Jersey). Samples were then heated to 90 degrees for 60 seconds and placed immediately on ice. Analysis was completed using the Megabace 1000 capillary electrophoresis system with Genetic Profiler Version 1.5 software (Amersham Biosciences, Piscataway, New Jersey). Injection parameters

were −3kV for 45 seconds and run parameters were −10kV for 75 minutes at 44°C.

The procedure to identify a fetal signal within that produced by MFPCR of the isolated cells was that outlined in the paper 'Analysis and interpretation of mixed forensic stains using DNA STR profiling' (Clayton *et al.*, 1998, Forensic Science International **91** 55-70). That is the STR is an additional band to that found in the maternal fingerprint i.e. consistent with maternal signal. It is not consistent with a stutter band or artefact peak and that it is the same base pair size as bands identified as fetal for the same locus within other isolations from the same patient.

Protocols for enrichment and diagnosis of fetal cells from the cervix must be consistently successful, robust and inexpensive if the techniques are to become an alternative to invasive procedures such as amniocentesis or chorionic villus sampling. Previous inventions and work has been significantly limited by a variety of factors including: obtaining sufficient cells; isolation of fetal cells from the sample; genetic identification of cells to determine fetal source; genetic diagnosis from small cell numbers and sample collection.

Where previous work has failed to detect and analyse fetal cells at a consistent and reliable enough level to be considered as a promising tool towards minimally invasive prenatal diagnosis, this invention provides incremental improvements to multiple steps and combines them into a high throughput method allowing widespread application for the first time.

However alternatively it should also be appreciated that this invention may also be considered as a complementary technique to other non-invasive or minimally invasive tests such as biochemical screening and

ultrasound screening offered to pregnant women during the first trimester of pregnancy (Daryani *et al.*, 2000, J. Obstet. Gynecol. **183** 752).

Our results indicate that cells reacting against multiple antibody sets are present in all patients. The percentage of fluorescent cells vary, with no apparent correlation to gestation. The overall number of fluorescent cells in the sample also varies with no correlation to gestation. This may indicate that variation in the number of fetal cells in the sample is specific to the patient or perhaps more likely due to variations in the technique used by the operator performing the retrieval.

DNA fingerprinting using MFPCR was used to confirm cell origin of the fluorescent cells from each antibody set and patient. This MFPCR technique has the advantage of being highly discriminating for cell origin even when applied to very close relatives such as mother and baby.

Combining the antibody and MFPCR data from all patients, it is possible to determine a minimum specificity for each antibody set towards fetal cells although this may be due to non-specific binding to non-cellular particles, binding to non-intact cells or strong binding to a maternal cell lineage.

6.7 STR Genetic analysis

Previous work using genetic diagnosis of limited numbers of fetal cells obtained from the uterine cervix using techniques such as FISH and PCR is very limited (methods are compared in Table 4). FISH analysis can only identify fetal cells if they are aneuploid or originate from a male fetus (Fejgin *et al.*, 2001, *supra*) – this is an important and considerable limitation to the use of such techniques for prenatal diagnosis. Other studies use PCR analysis to detect disorders however in most cases this is limited to the gene analysed and quantitative variations in the maternal and fetal

alleles, for example RH(D) analysis (Tutschek *et al.*, 1995, Prenatal Diagnosis **15** 951). Again this is an important and considerable limitation to the use of such techniques for prenatal diagnosis.

However MFPCR has the advantages of overcoming these limitations, as it is not limited by sex or individual gene alleles. MFPCR has an extremely high level of discrimination between closely related individuals, can be performed on single cells and provides multiple diagnoses within a single reaction.

In this embodiment, MFPCR was used to accurately determine the presence of fetal cells in a mixed fetal/maternal sample. For these reasons we suggest that MFPCR be considered the preferred method of choice when performing prenatal genetic diagnosis from pap smear samples.

Multiplex fluorescent PCR of Amelogenin and STRs D3S1358, D5S818, D7S820, CSF1P0, THO, FGA, D21S11, and D18S51 or any markers listed in Table 1, or more preferably Table 2, was performed on isolated cells using previously described protocols (Findlay *et al.*, 2001, *supra*; International Application PCT/AU02/01388). Each reaction contained forward and reverse primers, 1 X PCR buffer (Applied Biosystems, USA), 1.25mM each dNTP (Gibco, Life Technologies, Melbourne, Australia) and 1 unit Qiagen HotStarTaq (Qiagen Melbourne, Australia). PCR conditions were 94°C/2 minute denaturation followed by 45 cycles of 94°C/10 second denaturation, 57°C/1 minute annealing and 68°C/30 second extension. The PCR uses no oil overlay, as the heated lid of the PCR is sufficient. The PCR is taken off the block and stored at 4°C until required for electrophoresis.

PCR product was processed using Ammonium acetate/Ethanol Clean-up. Post clean-up processing involved adding 2uL of cleaned-up product to 3uL loading buffer (Amersham Biosciences, Piscataway, New Jersey).

Samples were then heated to 90 degrees for 60 seconds and placed immediately on ice. Analysis was completed using the Megabace 1000 capillary electrophoresis system with Genetic Profiler Version 1.5 software (Amersham Biosciences, Piscataway, New Jersey). Injection parameters were -3kV for 45 seconds and run parameters were -10kV for 75 minutes at 44°C.

The procedure to identify a fetal signal within that produced from MFPCR of the isolated cells was that outlined in Clayton *et al.*, 1998, *supra*.

This embodiment also indicates that cells of fetal origin are indeed present in the endocervical canal of the mother. Due to the relatively low number of fetal cells present serial enrichment strategies must be utilized to in unison to provide a reliable source of uncontaminated fetal cells yet still provide maximum yield. Once a reliable source of fetal cells is established, the diagnostic techniques such as MFPCR used to screen these cells needs to not only confirm fetal origin and but also test for genetic traits.

Again MFPCR has the advantage of being highly discriminating for cell origin even when applied to close relatives such as mother and baby. MFPCR can be performed on single cells and provides multiple diagnosis within a single reaction. In this embodiment MFPCR was used to accurately determine the presence of fetal cells isolated from a mixed fetal/maternal sample.

This embodiment demonstrates that samples highly enriched in fetal cells (>90%) can be produced even though an uncontaminated source of fetal cells from pap smears (i.e. isolation of 100% fetal cells) may not be possible. Single fetal cells can then be easily isolated and used to screen for genetic traits. For this reason, and the ability to test for multiple probes,

MFPCR may be considered the method of choice when performing prenatal genetic diagnosis from pap smear samples.

6.8 SNP genetic analysis

Many common diseases in humans are not caused by variation within single genes but are instead influenced by complex interactions among multiple genes as well as a multitude of environmental and lifestyle factors. Genetic factors may also confer susceptibility or resistance to a disease as well as determine the severity or progression of the disease. As most of the factors involved in these intricate pathways are unknown, it has therefore been difficult to develop screening tests for many diseases and disorders. It is therefore vital to understanding the genetic basis of common human diseases and this depends on a detailed understanding of the variability observed in the human genome.

Single Nucleotide Polymorphisms (SNP) are the most frequent form of variation found in the genome, estimated to occur every 1000 bases. SNP genotyping has multiple applications such as predictive medicine, personal medicine, forensic identification and pharmacogenomics However conventional SNP analysis is limited by the relatively high amount of extracted DNA usually required (up to 100ng) for analysis. However in genomic analysis, there is increasing demand to both maximize data by performing multiple analyses and secondly to analyze minimum amounts of sample, even to the single cell level. Although multiple SNP analyses can be performed routinely, the degree if sensitivity is still far from single cell level analysis. Multiplex single cell SNP analysis has been problematic and again

is not amenable to the high throughput processing required of clinical application.

SNP genotyping can be used to identify genetic regions associated with a disease phenotype, allowing researchers to target particular areas of interest and begin to reveal relevant genes associated with a disease. SNP patterns from a large group of affected individuals can be compared to those of unaffected individuals. These association studies can detect differences in the SNP patterns of the two groups, thereby indicating potentially important SNPs and thus genetic regions for further study. Eventually SNP profiles that are characteristic of a variety of diseases will become established. Defining and understanding the role of genetic factors in disease will also allow researchers to better evaluate the role that non-genetic factors - such as behaviour, diet, lifestyle, and physical activity - have on disease.

SNP genotyping has already been used to investigate a number of disorders such as cystic fibrosis, Factor V Leiden mutation, and factors such as A, B, O and Rh blood grouping.

SNP genotyping is undertaken in six main stages: PCR, Post-PCR cleanup, SNP primer extension reaction, final cleanup, SNP product sizing and analysis.

To provide an example of SNP genotyping we have utilised methods based around Amersham Megabace 1000 SNP manufacturers protocols.

6.8.1 PCR for SNPs

Isolated fetal cells were processes in a multiplex SNP reaction consisting of oligonucleotides for specific SNPs such as Kell, Rh etc. Each 25µl reaction contained 25pmol forward and reverse primers, 1 X PCR buffer, 5nM each dNTP (Gibco, Life Technologies, Melbourne, Australia) and 1 unit Qiagen HotStarTaq (Qiagen, Melbourne, Australia). PCR conditions were 95°C for 15 minute denaturation followed by 45 cycles of 20 secs at 94°C, 60°C then 72°C then followed by a two minute extension at 72°C.

6.8.2 Post PCR cleanup

Post PCR cleanup removes excess dNTPs and residual primers before primer extension and commonly uses SAP (shrimp alkaline phosphatase) and EXOI (exonuclease I) protocols.

Using a half-reaction protocol, 10µl of PCR product is added to 2.5µl EXO SAP (Amersham Biosciences) and incubated at 37°C for 15 minutes.

6.8.3 SNP primer extension reaction

Combine 1.0 µl SNuPe premix and 1µl dilution buffer (Amersham Biosciences) with 2 pmol extension primer, 1-10 ng cleaned PCR template (~1 ul) with distilled water to bring up to a final volume of 5µl. PCR conditions were 25 cycles of 96°C for 10 sec, 50° C for 5 sec then 60° C for 10 sec.

6.8.4 Final cleanup

The purpose of a final clean-up step is to remove excess terminators and desalts the samples prior to electrokinetic injection. AutoSeq96 columns (Amersham Biosciences) are used as per following protocol.

Spin product at 910g for 5 minutes, add 100µl deionised water, repeat spin, add samples to spin columns and repeat spin.

It is recommended to have at least one additional water wash to obtain sufficient yield good signal intensities. Additional water washes will further increase the signal intensity.

6.8.5 SNP product sizing

Add 2.5ul multiple injection marker (MIM, Amersham Biosciences) to 497.5ul loading solution then dispense 5µl into each well. Load LPA matrix (Amersham, Biosciences) and rerun as per manufacturers protocol. Perform cycles of sample injection then two-minute electrophoresis interval and repeat upto twelve times.

Enter sizing and SNP parameters into Snupe and Instrument Control Manager. Analysis is performed using SNP Profiler as per manufacturers protocols.

7 Overview

This specific example of the invention demonstrates that high throughput enrichment of fetal cells from pap smears can be performed for the first time using combination of improved methods including FACS, nucleic acid amplification, genetic analysis. The combinations of this technology with improved automated procedures for genetic identification

and analysis, have been applied to create a much improved method which allows automated high throughput system to maximize cost effectiveness and thus offer practical non-invasive prenatal analysis application for the first time.

This embodiment therefore represents a substantial advance compared to prior art and confirms that non-invasive prenatal diagnosis from pap smears can be automated to provide the high through capability required for clinical application.

Throughout this specification, the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. Various changes and modifications may be made to the embodiments described and illustrated herein without departing from the broad spirit and scope of the invention.

All patent and scientific literature, computer programs and algorithms referred to in this specification are incorporated herein by reference in their entirety.

7.1.1 TABLE 1 – STR markers used for DNA fingerprinting and genetic analysis

MARKER	ALIAS	GENBANK NO.	POSITION	PRIMER SEQUENCE
D13S241	UT556	L17673	13pter	CCA GGC ACT TTG GGA GGC TG ACC CAC TGT ATC CTG GGC A
D13S242	UT557	L18329	13q21.2	ATT GCA CCC CAT CCT GGG TCC TTT TCC TAC CAT TTG CAT
D13S243	UT558	L18330	13cen-13q12.1	ACT GTA CTT CTG CCT GGG C TTT TGT AAT GCC TCA ACC ATG
D13S248	UT1213	L15541	13q32-13q34	ACT TAA ATG TCC ATC AAT AAA T TGA TTG GCT TTT TTT ACT TAC
D13S251	UT1329	L16338	13q31-13q32	CAC ATA GCT TAT TGT TGT TGC GTT ATC TGT GAG CAA ATA CAG
D13S253	UT1378	L16396	13q22-13q32	CTC AAG GGA TGT TAA CAC AC AGG AGG AAA AAG TGG AGA AG
D13S254	UT1585	L18690	13q31-13q32	TGA ACT CCG GCC TGG GTG A TTT TGG AGC TGG GGA TGT C
D13S256	UT2120	L17977	13q14.1-13q22	CCT GGG CAA CAA GAG CAA A AGC AGA GAG ACA TAA TTG TG
D13S257	UT2119	L18729	13q14.1-13q21.1	CAA CAA GAG CAA AAC TCC AT AAG CAC ATA AGT TGG TAT GAA
D13S258	UT2413	L18095	13q21.2-13q31	ACC TGC CAA ATT TTA CCA GG GAC AGA GAG AGG GAA TAA ACC
D13S303	UT936	L31309	13q22-13q31	ACA TCG CTC CTT ACC CCA TC TGT ACC CAT TAA CCA TCC CCA
D13S631	UT7403	L18392	13q31-13q32	GGC AAC AAG AGC AAA ACT CT TAG CCC TCA CCA TGA TTG G
D18S51	UT574	L18333	18q21.33-18q21.33	GAG CCA TGT TCA TGC CAC TG CAA ACC CGA CTA CCA GCA AC

D18S378	UT485	L16262	18P11.22-18P11.22	AGC CTG GGT GAC AGA GCA A ACA GGG AAA GCT GGG GGA T
D18S382	UT600	L16292	18pter-18pter	CAT CCA TCC ATC CTT CCA C TGT GCT GGT ATT ACA GGC G
D18S386	UT754	L18400	18q22.1-18q22.2	TCA GGA GAA TCA CTT GGA AC TCC ATG AAG TAG CTA AGC AG
D18S390	UT1227	L15542	18q22.3-18q23	TAA CCA AAG CAA ATC CCT GG CAC TTA CAC TGT TAT CCT GG
D18S391	UT1302	L16384	18pter-18p11.22	CTG GTT TTC GTC TTG AGA AG CAC TAT TCC CAT CTG AGT CA
D18S814	UT1248	L17776	18pter-18pter	CTT CCC TGG GTA TCA AGA CT TCC CAC TAT ATG TAT GTT CAC C
D18S815	UT1438	L17819	18pter-18qter	GGC TGA GAC AGG AGA ATC AC CTC ACC AGG ATT TCC TTG C
D18S819	UT7251	L30411	18pter-18qter	ACC ACA GTT ACT AAG ATG TAA GCC TCC AGA AAA AAT TTC CA
D18S851	SHGC 4561	G08002	18pter-18qter	CTG TCC TCT AGG CTC ATT TAG C TTA TGA AGC AGT GAT GCC AA
D21S11	VS17T3	M84567	21q21-21q21	GTG AGT CAA TTC CCC AAG GTT GTA TTA GTC AAT GTT CTC C
D21S1240	UT656	L18360	21pter-21qter	GAG ACG GTA GGA AAA GGA G AGC CAA GTT CGA GCC ACT G
D21S1244	UT761	L16331	21q21-21q22.1	GTC CCC ATA TTG ATA AAC TAT T ATG AAT AGG GGA TAT GCT GG
D21S1413	UT7582	L30513	21pter-21pter	TTG CAG GGA AAC CAC AGT T TCC TTG GAA TAA ATT CCC GG
D21S1412	UT6930	L29680	21pter-21pter	CGG AGG TTG CAG TGA GTT G GGG AAG GCT ATG GAG GAG A
D21S1411	UT1355	L17803	21pter-21pter	ATG ATG AAT GCA TAG ATG GAT G AAT GTG TGT CCT TCC AGG C

PENTA E	PAUL1	AC027004	21q	
AMEL		M55418		CCC TGG GCT CTG TAA AGA ATA GTG ATC AGA GCT TAA ACT GGG AAG CTG
HUMTHO		D00269	11p15-15.5	GCT TCC GAG TGC AGG TCA CA CAG CTG CCC TAG TCA GCA C
TPOX		M68651	2p23-2pter	CAC TAG CAC CCA GAA CCG TC CCT TGT CAG CGT TTA TTT GCC
VWA		M25858	12p12-pter	CCC TAG TGG ATG ATA AGA ATA ATC AGT ATG GGA CAG ATG ATA AAT ACA TAG GAT GGA TGG
D3S1358		11449919	3p	ACT GCA GTC CAA TCT GGG T ATG AAA TCA ACA GAG GCT TG
D5S818		G08446	5q21-q31	GGG TGA TTT TCC TCT TTG GT TGA TTC CAA TCA TAG CCA CA
D7S820		G08616	7q	TGT CAT AGT TTA GAA CGA ACT AAC G CTG AGG TAT CAA AAA CTC AGA GG
CSF1PO	U63963	X14720	5q33.3-34	AAC CTG AGT CTG CCA AGG ACT AGC TTC CAC ACA CCA CTG GCC ATC TTC
FGA		M64982	4q28	GCC CCA TAG GTT TTG AAC TCA TGA TTT GTC TGT AAT TGC CAG C
D13S317		G09017	13q22-q31	ACA GAA GTC TGG GAT GTG GA GCC CAA AAA GAC AGA CAG AA
DYS14				CTT TCC ACA GCC ACA TTT GTC X CAT CCA GAG CGT CCC TGG CTT
D13S622	990			
D13S304	937			
D13S247	991			
D13S621	642			
D13S250	1250			
D13S633	7708			
D13S243	558			
D13S625	1587			

D13S246	740
D13S252	1352
D13S629	6870
D13S624	5236
D13S305	5177
D13S240	472
D13S249	1222
D13S257	2119
D13S256	2120
D13S626	5570
D13S242	557
D13S634	7875
D13S258	2413
D13S303	936
D13S921	2347
D13S251	1329
D13S628	6073
D13S253	1378
D13S627	5821
D13S631	7403
D13S254	1585
D13S248	1213
D18S999	7873
D18S820	7913
D18S818	7162
D18S391	1302
D18S378	485
D18S819	7251
D18S816	5780
D18S814	1248

D18S386	754			
D18S382	600			
D18S817	6365			
D18S815	1438			
D18S390	1227			
D18S812	5025			
D18S380	576			
D18S821	7934			
D21S1418	976			
D21S1411	1355			
D21S1241	657			
D21S1443	CHLC.GATA26A04			
D21S1240	656			
D21S1409	1305			
D21S1250	1332			
D21S1244	761			
D21S1249	1025			
D21S1245	762			
D21S1413	7582			
D21S1408	5040			
D21S1246	973			
D21S1412	6930			
DYS290	708			
DYS391	CHLC.GATA32C10			
D13S241	UT556	L17673	13pter	CCA GGC ACT TTG GGA GGC TG ACC CAC TGT ATC CTG GGC A
PD1		AP001752	21q	TGG AAG GTC GAA GCT GAA GTG A CCT GTG GCG TGT CTT TTT ACT TTC T
PE1		AC027004	15q	ATC ACT TGA ACC CAG GAG GTG GA GGG GAG GCT GTG TAA GAA GTG TT

D8S1179e		GO8710	8q24	TTT GGC CAG AAA CCT CTG TAG CC AAC TGA AAC CCT GTG CAT TGT TGT TG
CSF1POe		X14720	5q33	TCC AAC CTG AGT CTG CCA AGG A CTT CCA CAC ACC ACT GGC CAT CTT
D5S818e		G08446	5q21	ACA AGG GTG ATT TTC CTC TTT GGT ATC CCA AGT GAT TCC AAT CAT AGC CAC A
D7S820e		G08616	7q	TGT CAT AGT TTA GAA CGA ACT AAC GAT AG AAA TCT GAG GTA TCA AAA ACT CAG AGG
D21S11e		M84567	21p11.1	AAT ATG TGA GTC AAT TCC CCA AGT GAA T TGT ATT AGT CAA TGT TCT CCA GAG ACA
FGAe		M64982	4q28	TGC CCC ATA GGT TTT GAA CTC ACA GAG TGA TTT GTC TGT AAT TGC CAG C
TPOXe		M68651	2p23	CAC TAG CAC CCA GAA CCG TCG TGT CCT TGT CAG CGT TTA TTT GCC
THOe		D00269	11p15	GTG GGC TGA AAA GCT CCC GAT GTG ATT CCC ATT GGC CTG TTC CTC
D18S51e		AP001534	18q21.3	TGA GCC ATG TTC ATG CCA CTG ACA AAC CCG ACT ACC AGC AAC TT
D3S1358e		11449919	3p21	ACT GCA GTC CAA TCT GGG TGA CAG ATG AAA TCA ACA GAG GCT TGC ATG TAT C
BKMDY1			Y	GAA GTG CTC GGC ATT GTT AGG AT AGA TCC ATT TGC AGA CTG CCT TAT
BKMDY2			Y	AAG TGC TCG GCA TTG TTA GGA TT CTA AGC AGA TCC ATT TGC AGA CT
D13S317e		G09017	13q22-q31	CTT CCT ACC ACT GAA CAT AAA CTG CTT AA CAG TGA GCC AAG GTC GTG CCA
D13S258e	UT2413	L18095	13q21.2-13q31	ACC TGC CAA ATT TTA CCA GGA GGA GAC AGA GAG AGG GAA TAA ACC AAT AAG A
D13S631e	UT7403	L18392	13q31-13q32	GGC AAC AAG AGC AAA ACT CTG C

D18S391e	UT1302	L16384	18pter-18p11.22	TGG AAA AAT AAT TTC TGG GGG TGG GA CTG GTT TTC GTC TTG AGA AGT CAT G
D18S851e	SHGC 4561	G08002	18pter-18qter	CAC TAT TCC CAT CTG AGT CAC TCA G ACA CAC ACA AAC ATC TCT TTC TAT CTA TAT A
D21S1411e	UT1355	L17803	21pter-21pter	GCC TTT ATG AAG CAG TGA TGC CAA ATG ATG AAT GCA TAG ATG GAT GGA TG
D21S1412e	UT6930	L29680	21pter-21pter	AAT GTG TGT CCT TCC AGG CTT TCT CGG AGG TTG CAG TGA GTT GAG
D21S1413e	UT7582	L30513	21pter-21pter	GGG AAG GCT ATG GAG GAG A TTG CAG GGA AAC CAC AGT TAT ACA TTC
DYS14e	See TSPY			TCC TTG GAA TAA ATT CCC GGA AGT TTT CAT CCA GAG CGT CCC TGG C
D13S241	UT556	L17673	13pter	GCT TTC CAC AGC CAC ATT GGT CC CCA GGC ACT TTG GGA GGC TG
PD1		AP001752	21q	ACC CAC TGT ATC CTG GGC A TGG AAG GTC GAA GCT GAA GTG A
PE1		AC027004	15q	CCT GTG GCG TGT CTT TTT ACT TTC T ATC ACT TGA ACC CAG GAG GTG GA
D8S1179e		GO8710	8q24	GGG GAG GCT GTG TAA GAA GTG TT TTT GGC CAG AAA CCT CTG TAG CC
CSF1POe		X14720	5q33	AAC TGA AAC CCT GTG CAT TGT TGT TG TCC AAC CTG AGT CTG CCA AGG A
D5S818e		G08446	5q21	CTT CCA CAC ACC ACT GGC CAT CTT ACA AGG GTG ATT TTC CTC TTT GGT ATC
D7S820e		G08616	7q	CCA AGT GAT TCC AAT CAT AGC CAC A TGT CAT AGT TTA GAA CGA ACT AAC GAT AG
D21S11e		M84567	21p11.1	AAA TCT GAG GTA TCA AAA ACT CAG AGG AAT ATG TGA GTC AAT TCC CCA AGT GAA T TGT ATT AGT CAA TGT TCT CCA GAG ACA

FGAe		M64982	4q28	TGC CCC ATA GGT TTT GAA CTC ACA GAG TGA TTT GTC TGT AAT TGC CAG C
TPOXe		M68651	2p23	CAC TAG CAC CCA GAA CCG TCG TGT CCT TGT CAG CGT TTA TTT GCC
THOe		D00269	11p15	GTG GGC TGA AAA GCT CCC GAT GTG ATT CCC ATT GGC CTG TTC CTC
D18S51e		AP001534	18q21.3	TGA GCC ATG TTC ATG CCA CTG ACA AAC CCG ACT ACC AGC AAC TT
D3S1358e		11449919	3p21	ACT GCA GTC CAA TCT GGG TGA CAG ATG AAA TCA ACA GAG GCT TGC ATG TAT C
BKMDY1			Y	GAA GTG CTC GGC ATT GTT AGG AT AGA TCC ATT TGC AGA CTG CCT TAT
BKMDY2			Y	AAG TGC TCG GCA TTG TTA GGA TT CTA AGC AGA TCC ATT TGC AGA CT
D13S317e		G09017	13q22-q31	CTT CCT ACC ACT GAA CAT AAA CTG CTT AA CAG TGA GCC AAG GTC GTG CCA
D13S258e	UT2413	L18095	13q21.2-13q31	ACC TGC CAA ATT TTA CCA GGA GGA GAC AGA GAG AGG GAA TAA ACC AAT AAG A
D13S631e	UT7403	L18392	13q31-13q32	GGC AAC AAG AGC AAA ACT CTG C TGG AAA AAT AAT TTC TGG GGG TGG GA
D18S391e	UT1302	L16384	18pter-18p11.22	CTG GTT TTC GTC TTG AGA AGT CAT G CAC TAT TCC CAT CTG AGT CAC TCA G
D18S851e	SHGC 4561	G08002	18pter-18qter	ACA CAC ACA AAC ATC TCT TTC TAT CTA TAT A GCC TTT ATG AAG CAG TGA TGC CAA
D21S1411e	UT1355	L17803	21pter-21pter	ATG ATG AAT GCA TAG ATG GAT GGA TG AAT GTG TGT CCT TCC AGG CTT TCT
D21S1412e	UT6930	L29680	21pter-21pter	CGG AGG TTG CAG TGA GTT GAG GGG AAG GCT ATG GAG GAG A
D21S1413e	UT7582	L30513	21pter-21pter	TTG CAG GGA AAC CAC AGT TAT ACA TTC

DYS14e	See TSPY	TCC TTG GAA TAA ATT CCC GGA AGT TTT CAT CCA GAG CGT CCC TGG C GCT TTC CAC AGC CAC ATT GGT CC

7.1.2 TABLE 2 Example of Markers used for genetic analysis embodiment

Primer set	Fluorescent Dye	pmoles
Amelogenin	FAM	Variable from 1-40
DYS14	FAM	Variable from 1-40
D21S11	FAM or TET	Variable from 1-40
D13S631	HEX	Variable from 1-40
D13S258	HEX	Variable from 1-40
D18S51	FAM	Variable from 1-40
D18S851	FAM	Variable from 1-40
D18S391	HEX	Variable from 1-40
D13S317	TET	Variable from 1-40
D21S1413	HEX	Variable from 1-40
D21S1412	TET	Variable from 1-40
D21S1411	FAM	Variable from 1-40

7.1.3 TABLE 3 Example of Markers used for DNA fingerprinting embodiment

Primer set	Fluorescent Dye	pmoles
Amelogenin	FAM	Variable from 1-40
HUMTHO	FAM	Variable from 1-40
D21S11	FAM	Variable from 1-40
D18S51	FAM	Variable from 1-40
VWA	HEX	Variable from 1-40
FGA	HEX	Variable from 1-40
D3S1358	FAM	Variable from 1-40
D5S818	TET	Variable from 1-40
D7S820	TET	Variable from 1-40
CSF1PO	HEX	Variable from 1-40
TPOX	TET	Variable from 1-40

7.1.4 TABLE 4 Comparison of analysis methods

	Fluorescent PCR	FISH	Conventional PCR	PRINS
High reliability	97%	86%	84% for CF	91%
High accuracy	97-99% 97% for (Carrier) in CF	>95%	79% (Unaffect) & 66% (Carrier) for CF	25%
Rapid diagnosis	6 hrs	~4 hrs	8-10 hrs	6 hrs
Diagnosis of sex	Yes	Yes	Possible but poor reliability & accuracy	Yes
Diagnosis of single-gene defects	Yes	No	Yes	No
Diagnosis of trisomies	Yes	Yes	No	Yes
Confirmation of diagnosis	Yes	No	No	No
DNA fingerprinting	Yes, high specificity 1 in 100 million.	No	Limited ~1 in 10	No
Detection of contamination	Yes	No	Very limited	No
Simultaneous diagnoses	Sex, CF, trisomies & DNA fingerprint	Trisomies	No	No
No of chromosomes simultaneously analysed	~26	3-5	1	3